APR 0 9 2002

SEQUENCE LISTING

<110> Sem, Daniel S.
      Baker, Brian
      Hansen, Mark R.

<120> Methods for Predicting Functional and
   Structural Properties of Polypeptides Using Sequence Models

<130> P-TB 5072

<140> US 10/040,895
<141> 2001-12-28

<150> US 09/753,020
<151> 2000-12-29

<160> 17

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 155
<212> PRT
<213> Homo sapiens

<400> 1
Cys Leu Ile Gly Cys Gly Phe Ser Thr Gly Tyr Gly Ala Ala Val Lys
 1   5   10   15
Thr Gly Lys Val Lys Pro Gly Ser Thr Cys Val Val Phe Gly Leu Gly
            20   25   30
Gly Val Gly Leu Ser Val Ile Met Gly Cys Lys Ser Ala Gly Ala Ser
            35   40   45
Arg Ile Ile Gly Ile Asp Leu Asn Lys Asp Lys Phe Glu Lys Ala Met
     50   55   60
Ala Val Gly Ala Thr Glu Cys Ile Ser Pro Lys Asp Ser Thr Lys Pro
65   70   75   80
Ile Ser Glu Val Leu Ser Glu Met Thr Gly Asn Asn Val Gly Tyr Thr
                85   90   95
Phe Glu Val Ile Gly His Leu Glu Thr Met Ile Asp Ala Leu Ala Ser
            100   105   110
Cys His Met Asn Tyr Gly Thr Ser Val Val Val Gly Val Pro Pro Ser
            115   120   125
Ala Lys Met Leu Thr Tyr Asp Pro Met Leu Leu Phe Thr Gly Arg Thr
            130   135   140
Trp Lys Gly Cys Val Phe Gly Gly Leu Lys Ser
145   150   155

<210> 2
<211> 152

<212> PRT
<213> Equus caballus

<400> 2

```
Gly Cys Gly Phe Ser Thr Gly Tyr Gly Ser Ala Val Lys Val Ala Lys
1               5                   10                  15
Val Thr Gln Gly Ser Thr Cys Ala Val Phe Gly Leu Gly Gly Val Gly
            20                  25                  30
Leu Ser Val Ile Met Gly Cys Lys Ala Ala Gly Ala Ala Arg Ile Ile
            35                  40                  45
Gly Val Asp Ile Asn Lys Asp Lys Phe Ala Lys Ala Lys Glu Val Gly
        50                  55                  60
Ala Thr Glu Cys Val Asn Pro Gln Asp Tyr Lys Lys Pro Ile Gln Glu
65                  70                  75                  80
Val Leu Thr Glu Met Ser Asn Gly Gly Val Asp Phe Ser Phe Glu Val
                85                  90                  95
Ile Gly Arg Leu Asp Thr Met Val Thr Ala Leu Ser Cys Cys Gln Glu
            100                 105                 110
Ala Tyr Gly Val Ser Val Ile Val Gly Val Pro Pro Asp Ser Gln Asn
            115                 120                 125
Leu Ser Met Asn Pro Met Leu Leu Leu Ser Gly Arg Thr Trp Lys Gly
        130                 135                 140
Ala Ile Phe Gly Gly Phe Lys Ser
145                 150
```

<210> 3
<211> 175
<212> PRT
<213> Thermoanaerobium Brockii

<400> 3

```
Val Met Ile Pro Asp Met Met Thr Thr Gly Phe His Gly Ala Glu Leu
1               5                   10                  15
Ala Asp Ile Glu Leu Gly Ala Thr Val Ala Val Leu Gly Ile Gly Pro
            20                  25                  30
Val Gly Leu Met Ala Val Ala Gly Ala Lys Leu Arg Gly Ala Gly Arg
            35                  40                  45
Ile Ile Ala Val Gly Ser Arg Pro Val Cys Val Asp Ala Ala Lys Tyr
        50                  55                  60
Tyr Gly Ala Thr Asp Ile Val Asn Tyr Lys Asp Gly Pro Ile Glu Ser
65                  70                  75                  80
Gln Ile Met Asn Leu Thr Glu Gly Lys Gly Val Asp Ala Ala Ile Ile
                85                  90                  95
Ala Gly Gly Asn Ala Asp Ile Met Ala Thr Ala Val Lys Ile Val Lys
            100                 105                 110
Pro Gly Gly Thr Ile Ala Asn Val Asn Tyr Phe Gly Glu Gly Glu Val
            115                 120                 125
Leu Pro Val Pro Arg Leu Glu Trp Gly Cys Gly Met Ala His Lys Thr
        130                 135                 140
Ile Lys Gly Gly Leu Cys Pro Gly Gly Arg Leu Arg Met Glu Arg Leu
145                 150                 155                 160
Ile Asp Leu Val Phe Tyr Lys Arg Val Asp Pro Ser Lys Leu Val
```

165   170   175

<210> 4
<211> 141
<212> PRT
<213> Lactobacillus confusus

<400> 4
Ala Arg Lys Ile Gly Ile Ile Gly Leu Gly Asn Val Gly Ala Ala Val
1   5   10   15
Ala His Gly Leu Ile Ala Gln Gly Val Ala Asp Asp Tyr Val Phe Ile
        20   25   30
Asp Ala Asn Glu Ala Lys Val Lys Ala Asp Gln Ile Asp Phe Gln Asp
        35   40   45
Ala Met Ala Asn Leu Glu Ala His Gly Asn Ile Val Ile Asn Asp Trp
    50   55   60
Ala Ala Leu Ala Asp Ala Asp Val Val Ile Ser Thr Leu Gly Asn Ile
65   70   75   80
Lys Leu Gln Gln Phe Ala Glu Leu Lys Phe Thr Ser Ser Met Val Gln
            85   90   95
Ser Val Gly Thr Asn Leu Lys Glu Ser Gly Phe His Gly Val Leu Val
        100   105   110
Val Ile Ser Asn Pro Val Asp Val Ile Thr Ala Leu Phe Gln His Val
        115   120   125
Thr Gly Phe Pro Ala His Lys Val Ile Gly Thr Gly Thr
    130   135   140

<210> 5
<211> 147
<212> PRT
<213> B. Stearothermophilus

<400> 5
Met Lys Asn Asn Gly Gly Ala Arg Val Val Val Ile Gly Ala Gly Phe
1   5   10   15
Val Gly Ala Ser Tyr Val Phe Ala Leu Met Asn Gln Gly Ile Ala Asp
        20   25   30
Glu Ile Val Leu Ile Asp Ala Asn Glu Ser Lys Ala Ile Gly Asp Ala
        35   40   45
Met Asp Phe Asn His Gly Lys Val Phe Ala Pro Lys Pro Val Asp Ile
    50   55   60
Trp His Gly Asp Tyr Asp Asp Cys Arg Asp Ala Asp Leu Val Val Ile
65   70   75   80
Cys Ala Gly Ala Asn Gln Lys Pro Gly Glu Thr Arg Leu Asp Leu Val
            85   90   95
Asp Lys Asn Ile Ala Ile Phe Arg Ser Ile Val Glu Ser Val Met Ala
        100   105   110
Ser Gly Phe Gln Gly Leu Phe Leu Val Ala Thr Asn Pro Val Asp Ile
        115   120   125
Leu Thr Tyr Ala Thr Trp Lys Phe Ser Gly Leu Pro His Glu Arg Val
    130   135   140

Ile Gly Ser
145

<210> 6
<211> 312
<212> PRT
<213> E. Coli

<400> 6
Met Lys Val Ala Val Leu Gly Ala Ala Gly Gly Ile Gly Gln Ala Leu
 1   5   10   15
Ala Leu Leu Leu Lys Thr Gln Leu Pro Ser Gly Ser Glu Leu Ser Leu
            20   25   30
Tyr Asp Ile Ala Pro Val Thr Pro Gly Val Ala Val Asp Leu Ser His
            35   40   45
Ile Pro Thr Ala Val Lys Ile Lys Gly Phe Ser Gly Glu Asp Ala Thr
            50   55   60
Pro Ala Leu Glu Gly Ala Asp Val Val Leu Ile Ser Ala Gly Val Arg
65   70   75   80
Arg Lys Pro Gly Met Asp Arg Ser Asp Leu Phe Asn Val Asn Ala Gly
                85   90   95
Ile Val Lys Asn Leu Val Gln Gln Val Ala Lys Thr Cys Pro Lys Ala
            100   105   110
Cys Ile Gly Ile Ile Thr Asn Pro Val Asn Thr Thr Val Ala Ile Ala
            115   120   125
Ala Glu Val Leu Lys Lys Ala Gly Val Tyr Asp Lys Asn Lys Leu Phe
    130   135   140
Gly Val Thr Thr Leu Asp Ile Ile Arg Ser Asn Thr Phe Val Ala Glu
145   150   155   160
Leu Lys Gly Lys Gln Pro Gly Glu Val Glu Val Pro Val Ile Gly Gly
                165   170   175
His Ser Gly Val Thr Ile Leu Pro Leu Leu Ser Gln Val Pro Gly Val
                180   185   190
Ser Phe Thr Glu Gln Glu Val Ala Asp Leu Thr Lys Arg Ile Gln Asn
                195   200   205
Ala Gly Thr Glu Val Val Glu Ala Lys Ala Gly Gly Gly Ser Ala Thr
    210   215   220
Leu Ser Met Gly Gln Ala Ala Ala Arg Phe Gly Leu Ser Leu Val Arg
225   230   235   240
Ala Leu Gln Gly Glu Gln Gly Val Val Glu Cys Ala Tyr Val Glu Gly
                245   250   255
Asp Gly Gln Tyr Ala Arg Phe Phe Ser Gln Pro Leu Leu Leu Gly Lys
                260   265   270
Asn Gly Val Glu Glu Arg Lys Ser Ile Gly Thr Leu Ser Ala Phe Glu
                275   280   285
Gln Asn Ala Leu Glu Gly Met Leu Asp Thr Leu Lys Lys Asp Ile Ala
    290   295   300
Leu Gly Gln Glu Phe Val Asn Lys
305   310

<210> 7

<211> 163
<212> PRT
<213> Sus scrofa

<400> 7
```
Ala Thr Leu Lys Asp Gln Leu Ile His Asn Leu Leu Lys 'Glu Glu His
1               5                   10                  15
Val Pro His Asn Lys Ile Thr Val Val Gly Val Gly Ala Val Gly Met
            20                  25                  30
Ala Cys Ala Ile Ser Ile Leu Met Lys Glu Leu Ala Asp Glu Ile Ala
            35                  40                  45
Leu Val Asp Val Met Glu Asp Lys Leu Lys Gly Glu Met Met Asp Leu
        50                  55                  60
Gln His Gly Ser Leu Phe Leu Arg Thr Pro Lys Ile Val Ser Gly Lys
65                  70                  75                  80
Asp Tyr Asn Val Thr Ala Asn Ser Arg Leu Val Val Ile Thr Ala Gly
                85                  90                  95
Ala Arg Gln Gln Glu Gly Glu Ser Arg Leu Asn Leu Val Gln Arg Asn
            100                 105                 110
Val Asn Ile Phe Lys Phe Ile Ile Pro Asn Ile Val Lys Tyr Ser Pro
            115                 120                 125
Asn Cys Lys Leu Leu Val Val Ser Asn Pro Val Asp Ile Leu Thr Tyr
            130                 135                 140
Val Ala Trp Lys Ile Ser Gly Phe Pro Lys Asn Arg Val Ile Gly Ser
145                 150                 155                 160
Gly Cys Asn
```

<210> 8
<211> 333
<212> PRT
<213> Sus scrofa

<400> 8
```
Ser Glu Pro Ile Arg Val Leu Val Thr Gly Ala Ala Gly Gln Ile Ala
1               5                   10                  15
Tyr Ser Leu Leu Tyr Ser Ile Gly Asn Gly Ser Val Phe Gly Lys Asp
            20                  25                  30
Gln Pro Ile Ile Leu Val Leu Leu Asp Ile Thr Pro Met Met Gly Val
            35                  40                  45
Leu Asp Gly Val Leu Met Glu Leu Gln Asp Cys Ala Leu Pro Leu Leu
        50                  55                  60
Lys Asp Val Ile Ala Thr Asp Lys Glu Glu Ile Ala Phe Lys Asp Leu
65                  70                  75                  80
Asp Val Ala Ile Leu Val Gly Ser Met Pro Arg Arg Asp Gly Met Glu
                85                  90                  95
Arg Lys Asp Leu Leu Lys Ala Asn Val Lys Ile Phe Lys Cys Gln Gly
            100                 105                 110
Ala Ala Leu Asp Lys Tyr Ala Lys Lys Ser Val Lys Val Ile Val Val
            115                 120                 125
Gly Asn Pro Ala Asn Thr Asn Cys Leu Thr Ala Ser Lys Ser Ala Pro
            130                 135                 140
```

```
Ser Ile Pro Lys Glu Asn Phe Ser Cys Leu Thr Arg Leu Asp His Asn
145                 150                 155                 160
Arg Ala Lys Ala Gln Ile Ala Leu Lys Leu Gly Val Thr Ser Asp Asp
                165                 170                 175
Val Lys Asn Val Ile Ile Trp Gly Asn His Ser Ser Thr Gln Tyr Pro
             ·  180                 185                 190
Asp Val Asn His Ala Lys Val Lys Leu Gln Ala Lys Glu Val Gly Val
            195                 200                 205
Tyr Glu Ala Val Lys Asp Asp Ser Trp Leu Lys Gly Glu Phe Ile Thr
    210                 215                 220
Thr Val Gln Gln Arg Gly Ala Ala Val Ile Lys Ala Arg Lys Leu Ser
225                 230                 235                 240
Ser Ala Met Ser Ala Ala Lys Ala Ile Cys Asp His Val Arg Asp Ile
                245                 250                 255
Trp Phe Gly Thr Pro Glu Gly Glu Phe Val Ser Met Gly Ile Ile Ser
                260                 265                 270
Asp Gly Asn Ser Tyr Gly Val Pro Asp Asp Leu Leu Tyr Ser Phe Pro
            275                 280                 285
Val Thr Ile Lys Asp Lys Thr Trp Lys Ile Val Glu Gly Leu Pro Ile
            290                 295                 300
Asn Asp Phe Ser Arg Glu Lys Met Asp Leu Thr Ala Lys Glu Leu Ala
305                 310                 315                 320
Glu Glu Lys Glu Thr Ala Phe Glu Phe Leu Ser Ser Ala
                325                 330
```

<210> 9
<211> 159
<212> PRT
<213> Thermus Flavis

<400> 9
```
Met Lys Ala Pro Val Arg Val Ala Val Thr Gly Ala Ala Gly Gln Ile
1               5                   10                  15
Gly Tyr Ser Leu Leu Phe Arg Ile Ala Ala Gly Glu Met Leu Gly Lys
                20                  25                  30
Asp Gln Pro Val Ile Leu Gln Leu Leu Glu Ile Pro Gln Ala Met Lys
            35                  40                  45
Ala Leu Glu Gly Val Val Met Glu Leu Glu Asp Cys Ala Phe Pro Leu
            50                  55                  60
Leu Ala Gly Leu Glu Ala Thr Asp Asp Pro Asp Val Ala Phe Lys Asp
65                  70                  75                  80
Ala Asp Tyr Ala Leu Leu Val Gly Ala Ala Pro Arg Lys Ala Gly Met
                85                  90                  95
Glu Arg Arg Asp Leu Leu Gln Val Asn Gly Lys Ile Phe Thr Glu Gln
                100                 105                 110
Gly Arg Ala Leu Ala Glu Val Ala Lys Lys Asp Val Lys Val Leu Val
            115                 120                 125
Val Gly Asn Pro Ala Asn Thr Asn Ala Leu Ile Ala Tyr Lys Asn Ala
            130                 135                 140
Pro Gly Leu Asn Pro Arg Asn Phe Thr Ala Met Thr Arg Leu Asp
145                 150                 155
```

```
<210> 10
<211> 200
<212> PRT
<213> E. coli

<400> 10
Pro Phe Ser Asn Thr Arg Ser Val Ala Glu Leu Val Ile Gly Glu Leu
1               5                   10                  15
Leu Leu Leu Leu Arg Gly Val Pro Glu Ala Asn Ala Lys Ala His Arg
            20                  25                  30
Gly Val Trp Asn Lys Leu Ala Ala Gly Ser Phe Glu Ala Arg Gly Lys
            35                  40                  45
Lys Leu Gly Ile Ile Gly Tyr Gly His Ile Gly Thr Gln Leu Gly Ile
        50                  55                  60
Leu Ala Glu Ser Leu Gly Met Tyr Val Tyr Phe Tyr Asp Ile Glu Asn
65                  70                  75                  80
Lys Leu Pro Leu Gly Asn Ala Thr Gln Val Gln His Leu Ser Asp Leu
                85                  90                  95
Leu Asn Met Ser Asp Val Val Ser Leu His Val Pro Glu Asn Pro Ser
            100                 105                 110
Thr Lys Asn Met Met Gly Ala Lys Glu Ile Ser Leu Met Lys Pro Gly
            115                 120                 125
Ser Leu Leu Ile Asn Ala Ser Arg Gly Thr Val Val Asp Ile Pro Ala
    130                 135                 140
Leu Cys Asp Ala Leu Ala Ser Lys His Leu Ala Gly Ala Ala Ile Asp
145                 150                 155                 160
Val Phe Pro Thr Glu Pro Ala Thr Asn Ser Asp Pro Phe Thr Ser Pro
                165                 170                 175
Leu Cys Glu Phe Asp Asn Val Leu Leu Thr Pro His Ile Gly Gly Ser
                180                 185                 190
Thr Gln Glu Ala Gln Glu Asn Ile
                195                 200


<210> 11
<211> 236
<212> PRT
<213> L. casei

<400> 11
Ser Asn Val Pro Ala Tyr Ser Pro Ala Ala Ile Ala Glu Phe Ala Leu
1               5                   10                  15
Thr Asp Thr Leu Tyr Leu Leu Arg Asn Met Gly Lys Val Gln Ala Gln
            20                  25                  30
Leu Gln Ala Gly Asp Tyr Glu Lys Ala Gly Thr Phe Ile Gly Lys Glu
            35                  40                  45
Leu Gly Gln Gln Thr Val Gly Val Met Gly Thr Gly His Ile Gly Gln
        50                  55                  60
Val Ala Ile Lys Leu Phe Lys Gly Phe Gly Ala Lys Val Ile Ala Tyr
65                  70                  75                  80
Asp Pro Tyr Pro Met Lys Gly Asp His Pro Asp Phe Asp Tyr Val Ser
                85                  90                  95
```

```
Leu Glu Asp Leu Phe Lys Gln Ser Asp Val Ile Asp Leu His Val Pro
            100                 105                 110
Gly Ile Glu Gln Asn Thr His Ile Ile Asn Glu Ala Ala Phe Asn Leu
            115                 120                 125
Met Lys Pro Gly Ala Ile Val Ile Asn Thr Ala Arg Pro Asn Leu Ile
            130      .          135                 140           .
Asp Thr Gln Ala Met Leu Ser Asn Leu Lys Ser Gly Lys Leu Ala Gly
145                 150                 155                 160
Val Gly Ile Asp Thr Tyr Glu Tyr Glu Thr Glu Asp Leu Leu Asn Leu
                165                 170                 175
Ala Lys His Gly Ser Phe Lys Asp Pro Leu Trp Asp Glu Leu Leu Gly
            180                 185                 190
Met Pro Asn Val Val Leu Ser Pro His Ile Ala Tyr Tyr Thr Glu Thr
            195                 200                 205
Ala Val His Asn Met Val Tyr Phe Ser Leu Gln His Leu Val Asp Phe
            210                 215                 220
Leu Thr Lys Gly Glu Thr Ser Thr Glu Val Thr Gly
225                 230                 235


<210> 12
<211> 192
<212> PRT
<213> Methylotrophic bacterium pseudomonas sp.

<400> 12
Cys Asn Ser Ile Ser Val Ala Glu His Val Val Met Met Ile Leu Ser
 1               5                   10                  15
Leu Val Arg Asn Tyr Leu Pro Ser His Glu Trp Ala Arg Lys Gly Gly
            20                  25                  30
Trp Asn Ile Ala Asp Cys Val Ser His Ala Tyr Asp Leu Glu Ala Met
            35                  40                  45
His Val Gly Thr Val Ala Ala Gly Arg Ile Gly Leu Ala Val Leu Arg
        50                  55                  60
Arg Leu Ala Pro Phe Asp Val His Leu His Tyr Thr Asp Arg His Arg
65                  70                  75                  80
Leu Pro Glu Ser Val Glu Lys Glu Leu Asn Leu Thr Trp His Ala Thr
                85                  90                  95
Arg Glu Asp Met Tyr Pro Val Cys Asp Val Val Thr Leu Asn Cys Pro
            100                 105                 110
Leu His Pro Glu Thr Glu His Met Ile Asn Asp Glu Thr Leu Lys Leu
            115                 120                 125
Phe Lys Arg Gly Ala Tyr Ile Val Asn Thr Ala Arg Gly Lys Leu Cys
            130                 135                 140
Asp Arg Asp Ala Val Ala Arg Ala Leu Glu Ser Gly Arg Leu Ala Gly
145                 150                 155                 160
Tyr Ala Gly Asp Val Trp Phe Pro Gln Pro Ala Pro Lys Asp His Pro
                165                 170                 175
Trp Arg Thr Met Pro Tyr Asn Gly Met Thr Pro His Ile Ser Gly Thr
            180                 185                 190


<210> 13
```

<211> 131
<212> PRT
<213> Homo sapiens

<400> 13
Pro Cys Thr Pro Lys Gly Cys Leu Glu Leu Ile Lys Glu Thr Gly Val
1   5   10   15
Pro Ile Ala Gly Arg His Ala Val Val Val Gly Arg Ser Lys Ile Val
            20   25   30
Gly Ala Pro Met His Asp Leu Leu Leu Trp Asn Asn Ala Thr Val Thr
            35   40   45
Thr Cys His Ser Lys Thr Ala His Leu Asp Glu Glu Val Asn Lys Gly
    50   55   60
Asp Ile Leu Val Val Ala Thr Gly Gln Pro Glu Met Val Lys Gly Glu
65   70   75   80
Trp Ile Lys Pro Gly Ala Ile Val Ile Asp Cys Gly Ile Asn Tyr Lys
                85   90   95
Val Val Gly Asp Val Ala Tyr Asp Glu Ala Lys Glu Arg Ala Ser Phe
            100   105   110
Ile Thr Pro Val Pro Gly Gly Val Gly Pro Met Thr Val Ala Met Leu
            115   120   125
Met Gln Ser
    130

<210> 14
<211> 170
<212> PRT
<213> Rattus sp.

<400> 14
Lys Phe Asp Asn Leu Tyr Gly Cys Arg Glu Ser Leu Ile Asp Gly Ile
1   5   10   15
Lys Arg Ala Thr Asp Val Met Ile Ala Gly Lys Val Ala Val Val Ala
            20   25   30
Gly Tyr Gly Asp Val Gly Lys Gly Cys Ala Gln Ala Leu Arg Gly Phe
            35   40   45
Gly Ala Arg Val Ile Ile Thr Glu Ile Asp Pro Ile Asn Ala Leu Gln
    50   55   60
Ala Ala Met Glu Gly Tyr Glu Val Thr Thr Met Asp Glu Ala Cys Lys
65   70   75   80
Glu Gly Asn Ile Phe Val Thr Thr Thr Gly Cys Val Asp Ile Ile Leu
                85   90   95
Gly Arg His Phe Glu Gln Met Lys Asp Asp Ala Ile Val Cys Asn Ile
            100   105   110
Gly His Phe Asp Val Glu Ile Asp Val Lys Trp Leu Asn Glu Asn Ala
            115   120   125
Val Glu Lys Val Asn Ile Lys Pro Gln Val Asp Arg Tyr Leu Leu Lys
    130   135   140
Asn Gly His Arg Ile Ile Leu Leu Ala Glu Gly Arg Leu Val Asn Leu
145   150   155   160
Gly Cys Ala Met Gly His Pro Ser Phe Val
            165   170

<210> 15
<211> 179
<212> PRT
<213> Phormidium lapideum

<400> 15
Thr Pro Met Ser Ile Ile Ala Gly Arg Leu Ser Val Gln Phe Gly Ala
1   5   10   15
Arg Phe Leu Glu Arg Gln Gln Gly Gly Arg Gly Val Leu Leu Gly Gly
            20   25   30
Val Pro Gly Val Lys Pro Gly Lys Val Val Ile Leu Gly Gly Gly Val
            35   40   45
Val Gly Thr Glu Ala Ala Lys Met Ala Val Gly Leu Gly Ala Gln Val
            50   55   60
Gln Ile Phe Asp Ile Asn Val Glu Arg Leu Ser Tyr Leu Glu Thr Leu
65   70   75   80
Phe Gly Ser Arg Val Glu Leu Leu Tyr Ser Asn Ser Ala Glu Ile Glu
                85   90   95
Thr Ala Val Ala Glu Ala Asp Leu Leu Ile Gly Ala Val Leu Val Pro
                100   105   110
Gly Arg Arg Ala Pro Ile Leu Val Pro Ala Ser Leu Val Glu Gln Met
            115   120   125
Arg Thr Gly Ser Val Ile Val Asp Val Ala Val Asp Gln Gly Gly Cys
            130   135   140
Val Glu Thr Leu His Pro Thr Ser His Thr Gln Pro Thr Tyr Glu Val
145   150   155   160
Phe Gly Val Val His Tyr Gly Val Pro Asn Met Pro Gly Ala Val Pro
                165   170   175
Trp Thr Ala

<210> 16
<211> 165
<212> PRT
<213> E. coli

<400> 16
Asn Ile Arg Val Ala Ile Ala Gly Ala Gly Gly Arg Met Gly Arg Gln
1   5   10   15
Leu Ile Gln Ala Ala Leu Ala Leu Glu Gly Val Gln Leu Gly Ala Ala
            20   25   30
Leu Glu Arg Glu Gly Ser Ser Leu Leu Gly Ser Asp Ala Gly Glu Leu
            35   40   45
Ala Gly Ala Gly Lys Thr Gly Val Thr Val Gln Ser Ser Leu Asp Ala
        50   55   60
Val Lys Asp Asp Phe Asp Val Phe Ile Asp Phe Thr Arg Pro Glu Gly
65   70   75   80
Thr Leu Asn His Leu Ala Phe Cys Arg Gln His Gly Lys Gly Met Val
                85   90   95
Ile Gly Thr Thr Gly Phe Asp Glu Ala Gly Lys Gln Ala Ile Arg Asp

```
                100                    105                    110
Ala Ala Ala Asp Ile Ala Ile Val Phe Ala Ala Asn Phe Ser Val Gly
          115                    120                    125
Ala Ser Ser Arg Met Thr Phe Ala Asn Gly Ala Val Arg Ser Ala Leu
      130                    135                    140
Trp Leu Ser Gly Lys Glu Ser Gly Leu Phe Asp Met Arg Asp Val Leu
145                    150                    155                    160
Asp Leu Asn Asn Leu
                165
```

<210> 17
<211> 301
<212> PRT
<213> Homo sapiens

<400> 17

```
Leu Ile Gln Phe Glu Asp Phe Gly Asn His Asn Ala Phe Arg Phe Leu
  1               5                  10                     15
Arg Lys Tyr Arg Glu Lys Tyr Cys Thr Phe Asn Asp Asp Ile Gln Gly
          20                  25                     30
Thr Ala Ala Val Ala Leu Ala Gly Leu Leu Ala Ala Gln Lys Val Ile
          35                  40                     45
Ser Lys Pro Ile Ser Glu His Lys Ile Leu Phe Leu Gly Ala Gly Glu
      50                  55                     60
Ala Ala Leu Gly Ile Ala Asn Leu Ile Val Ser Val Glu Asn Gly Leu
65                  70                  75                     80
Ser Glu Gln Glu Ala Gln Lys Lys Ile Trp Phe Asp Lys Tyr Gly Leu
              85                  90                     95
Leu Val Lys Gly Arg Lys Ala Lys Ile Asp Ser Tyr Gln Glu Pro Phe
          100                    105                    110
Thr His Ser Ala Pro Glu Ser Ile Pro Asp Thr Phe Glu Asp Ala Val
          115                    120                    125
Asn Ile Leu Lys Pro Ser Thr Ile Ile Gly Val Ala Gly Ala Gly Arg
      130                    135                    140
Leu Phe Thr Pro Asp Val Ile Arg Ala Ala Ser Ile Asn Glu Arg Pro
145                    150                    155                    160
Val Ile Phe Ala Leu Ser Asn Pro Thr Ala Gln Ala Glu Cys Thr Ala
                  165                    170                    175
Glu Glu Ala Tyr Thr Leu Thr Glu Gly Arg Cys Leu Phe Ala Ser Gly
              180                    185                    190
Ser Pro Phe Gly Pro Val Lys Leu Thr Asp Gly Arg Val Phe Thr Pro
          195                    200                    205
Gly Gln Gly Asn Asn Val Tyr Ile Phe Pro Gly Val Ala Leu Ala Val
      210                    215                    220
Ile Leu Cys Asn Thr Arg His Ile Ser Asp Ser Val Phe Leu Glu Ala
225                    230                    235                    240
Ala Lys Ala Leu Thr Ser Gln Leu Thr Asp Glu Glu Leu Ala Gln Gly
                  245                    250                    255
Arg Leu Tyr Pro Pro Leu Ala Asn Ile Gln Glu Val Ser Ile Asn Ile
              260                    265                    270
Ala Ile Lys Val Thr Glu Tyr Leu Tyr Ala Asn Lys Ala Phe Arg Tyr
              275                    280                    285
```

Pro Glu Pro Glu Asp Lys Ala Lys Tyr Val Lys Glu Arg
   290   295   300